

Schriftliche Prüfungen: Vorteile von Multiple True-False Fragen gegenüber Typ-A

Autoren: Lahner, F.-M., Nouns, Z., Krebs, R., Fischer, M. & Huwendiek, S.

Fragestellung/Zielsetzung

Prüfungen sind essentieller Bestandteil in der ärztlichen Ausbildung. Sie liefern wertvolle Informationen über den Entwicklungsprozess der Studierenden und wirken lernbegleitend und lernmodulierend [1, 2].

Bei schriftlichen Prüfungen dominieren Multiple Choice Fragen, die in verschiedenen Typen verwendet werden. Zumeist werden Typ-A Fragen genutzt, bei denen genau eine Antwort richtig ist. Multiple True-False (MTF) Fragen hingegen lassen mehrere richtige Antworten zu: es muss für jede Antwortmöglichkeit entschieden werden, ob diese richtig oder falsch ist. Durch die Mehrfachantwort scheinen MTF Fragen bestimmte klinische Sachverhalte besser widerspiegeln zu können. Auch bezüglich Reliabilität und dem Informationsgewinn pro Testzeit scheinen MTF Fragen den Typ-A Fragen überlegen zu sein [3]. Dennoch werden MTF Fragen bislang selten genutzt und es gibt wenig Literatur zu diesem Fragenformat.

In dieser ersten Studie eines PhD-Forschungsprojekts soll untersucht werden, inwiefern die Verwendung von MTF Fragen die Nutzbarkeit (Utility) nach van der Vleuten (Reliabilität, Validität, Kostenaufwand, Effekt auf den Lernprozess und Akzeptanz der Teilnehmer) [4] schriftlicher Prüfungen erhöhen kann. Um die Testreliabilität zu steigern, sowie den Kostenaufwand für Prüfungen zu senken, möchten wir das optimale Bewertungssystem (Scoring) für MTF Fragen ermitteln.

Methoden

Wir analysieren die Daten summativer Prüfungen der medizinischen Fakultät der Universität Bern. Unsere Daten beinhalten Prüfungen vom ersten bis zum sechsten Studienjahr, sowie eine Facharztprüfung. Alle Prüfungen umfassen sowohl MTF als auch Typ-A Fragen.

Für diese Prüfungen vergleichen wir die Viertel-, Halb- und Ganzpunktbewertung für MTF Fragen. Bei der Viertelpunktbewertung bekommen Kandidaten für jede richtige Teilantwort $\frac{1}{4}$ Punkt. Bei der Halbpunktbewertung wird $\frac{1}{2}$ Punkt vergeben, wenn mehr als die Hälfte der Antwortmöglichkeiten richtig ist, einen ganzen Punkt erhalten die Kandidaten wenn alle Antworten richtig beantwortet wurden. Bei der Ganzpunktbewertung erhalten Kandidaten lediglich einen Punkt wenn die komplette Frage richtig beantwortet wurde. Diese unterschiedlichen Bewertungsschemata werden hinsichtlich Fragencharakteristika wie Trennschärfe und Schwierigkeit sowie hinsichtlich Testcharakteristika wie der Reliabilität einander gegenübergestellt. Die Ergebnisse werden ausserdem mit denen für Typ A Fragen verglichen.

Ergebnisse

Vorläufige Ergebnisse deuten darauf hin, dass eine Halbpunktbewertung optimal zu sein scheint. Eine Halbpunktbewertung führt zu mittleren Item-Schwierigkeiten und daraus resultierend zu hohen Trennschärfen. Dies trägt zu einer hohen Testreliabilität bei.

Diskussion/Schlussfolgerung

MTF Fragen scheinen in Verbindung mit einem optimalen Bewertungssystem, zu höheren Testreliabilitäten im Vergleich zu Typ A Fragen zu führen. In Abhängigkeit des zu prüfenden Inhalts könnten MTF Fragen eine wertvolle Ergänzung zu Typ-A Fragen darstellen. Durch die geeignete Kombination von MTF und Typ A Fragen könnte die Nutzbarkeit (Utility) schriftlicher Prüfungen verbessert werden.

1. Cox, M., D.M. Irby, and R.M. Epstein, *Assessment in medical education*. New England Journal of Medicine, 2007. **356**(4): p. 387-396.
2. De Champlain, A.F., *A primer on classical test theory and item response theory for assessments in medical education*. Medical education, 2010. **44**(1): p. 109-117.
3. Kreiter, C.D. and D.A. Frisbie, *Effectiveness of multiple true-false items*. Applied Measurement in Education, 1989. **2**(3): p. 207-216.
4. Vleuten, C.v., *The assessment of professional competence: developments, research and practical implications*. Advances in Health Sciences Education, 1996. **1**(1): p. 41-67.